

EmoCon: Analizador de Emociones en el Congreso de los Diputados

EmoCon: Emotions Analyzer in the Spanish Congress

Salud María Jiménez-Zafra, Miguel Ángel García-Cumbreras,
María Teresa Martín-Valdivia, Andrea López-Fernández
SINAI, Centro de Estudios Avanzados en TIC (CEATIC), Universidad de Jaén
{sjzafra, magc, maite, alfernan}@ujaen.es

Resumen: EmoCon es un prototipo de un analizador de emociones en el Congreso de los Diputados. Su objetivo es analizar el perfil emocional a nivel de sesión parlamentaria y a nivel de cada diputado, a partir de las intervenciones realizadas durante las sesiones parlamentarias que tienen lugar en el Congreso de los Diputados. Para ello, la demo cuenta con tres módulos principales: i) descarga automática de los documentos de las sesiones y extracción de las intervenciones realizadas por cada diputado, ii) análisis de las emociones expresadas a nivel de sesión y a nivel de diputado y, iii) visualización de la información en una aplicación web.

Palabras clave: Análisis de emociones, Congreso de los Diputados, política

Abstract: EmoCon is a prototype of an emotion analyzer in the Spanish Congress. Its objective is to analyze the emotions expressed by the deputies in the interventions made during the parliamentary sessions that take place in the Spanish Congress. To this end, the demo has three main modules: i) web scrapper for the session documents and processing, ii) emotions analyzer at the session level and at the deputy level, and iii) web application for visualization.

Keywords: Emotion analysis, Spanish Congress, politics

1 Introducción

Las emociones son un aspecto muy importante de nuestra vida. Lo que hacemos y lo que decimos refleja en parte las emociones que sentimos. La RAE define emoción como una “alteración del ánimo intensa y pasajera, agradable o penosa, que va acompañada de cierta conmoción somática”. Desde el punto de vista psicológico las emociones se suelen definir como un complejo estado afectivo, una reacción subjetiva que ocurre como resultado de cambios fisiológicos o psicológicos que influyen sobre el pensamiento y la conducta. Estas emociones se pueden analizar a través de las expresiones faciales, de la voz, del texto, etc.

La detección de emociones ha sido un tema de gran interés en disciplinas como la neurociencia o la psicología y, actualmente, está atrayendo la atención de los investigadores en ciencias de la computación. El área que se encarga del estudio de las emociones a nivel computacional se conoce como *Computación Afectiva* (Cambria, 2016). En este trabajo nos centramos en esta área, en la ta-

rea de reconocimiento de emociones en textos, que consiste en identificar de forma automática las emociones expresadas en un texto. En concreto, presentamos una demo que tiene como objetivo aplicar técnicas y herramientas de las Tecnologías del Lenguaje Humano para analizar y detectar las emociones que expresan los políticos en los diarios de sesiones del Congreso de los Diputados, fruto de sus intervenciones. La finalidad de esta demo es mostrar el perfil emocional de los diputados del Congreso a partir de sus intervenciones.

Existen varias clasificaciones posibles de categorías de emociones (Ekman, 1992; Plutchik, 1986; Lövhheim, 2012). EmoCon utiliza la clasificación de Paul Ekman (Ekman, 1992), al ser una de las más usadas y que más recursos tiene disponibles. Ekman define seis emociones básicas: enfado, miedo, asco, sorpresa, alegría y tristeza.

2 Motivación

En el ámbito de la política, el análisis de emociones se ha utilizado fundamentalmente para saber qué actitudes y propuestas de los candi-

datos electorales provocan en los ciudadanos mejores respuestas emocionales, con el objetivo de realizar sondeos políticos en vista a unas elecciones. Nuestro enfoque es bien distinto y está centrado en el análisis de las emociones de los propios parlamentarios por sus propios contenidos, los cuales podrían estar relacionados con la reacción de los ciudadanos.

3 Descripción del sistema

EmoCon extrae y analiza los diarios de sesión de las intervenciones de los diputados en el Congreso de los Diputados (debate político), detectando las emociones en los mismos. Cuenta con tres módulos principales:

1. Descarga de documentos de las sesiones y procesamiento
2. Analizador automático de emociones
3. Aplicación web para la visualización de la información procesada

La Figura 1 muestra el esquema general de EmoCon.

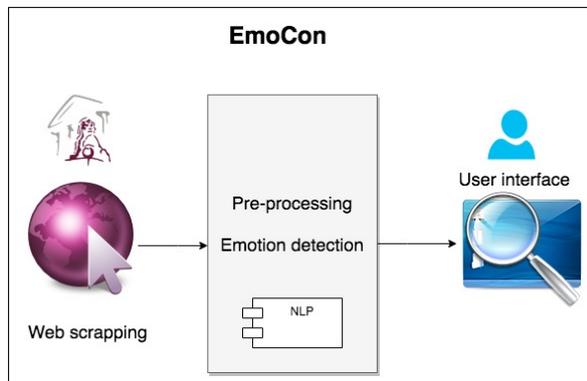


Figura 1: Esquema general de EmoCon

A continuación se detallan cada uno de estos módulos.

3.1 Módulo 1: descarga y procesamiento de datos

El sistema trabaja con información de las intervenciones realizadas en las sesiones parlamentarias que tienen lugar en el Congreso de los Diputados. Estos documentos están accesibles a través del portal web del Congreso de los Diputados¹. Los documentos están clasificados por legislatura, año, y mes, lo que facilita la localización de una sesión específica.

¹<http://www.congreso.es/portal/page/portal/Congreso/Congreso/Publicaciones/DiaSes/Pleno>

En concreto, la demo presentada realiza un análisis de la XII legislatura, es decir la que transcurre desde el 19 de julio de 2016 hasta el 5 de marzo de 2019, y se continúan descargando y analizando documentos de la legislatura vigente. Actualmente el sistema trabaja con un total de 640 documentos de sesión de 384 diputados.

Para la descarga de la información se ha implementado un crawler, aprovechando que las URLs de estos documentos siguen un patrón², y se ha utilizado Python como lenguaje de programación. En la Figura 2 se muestra un ejemplo de un extracto de uno de los PDFs descargados.

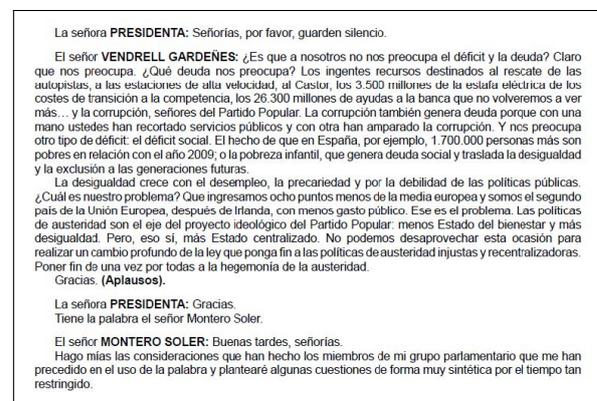


Figura 2: Extracto de un PDF de una sesión

De cada uno de los PDFs descargados el sistema extrae las intervenciones de los diputados, texto que posteriormente es analizado en el módulo 2 para determinar las emociones expresadas por cada diputado. Esto es posible ya que como se puede observar en la Figura 2, en estos documentos se sigue un patrón para mostrar cada una de las intervenciones de los distintos diputados: *El señor/La señora + nombre en mayúscula y negrita + dos puntos + intervención del diputado.*

3.2 Módulo 2: análisis de emociones

Este módulo es el encargado de analizar las emociones presentes en los textos extraídos de cada una de las sesiones parlamentarias. El sistema realiza dos tipos de análisis: uno a nivel de sesión y otro a nivel de diputado.

Para analizar las emociones utiliza el modelo de las seis emociones de Ekman comentado previamente y un clasificador automático basado en lexicón. El clasificador hace uso

²[http://www.congreso.es/public_oficiales/L12/CONG/DS/PL/DSCD-12-PL-\[numsesion\].PDF](http://www.congreso.es/public_oficiales/L12/CONG/DS/PL/DSCD-12-PL-[numsesion].PDF)

del lexicón “Spanish Emotion Lexicon”, conocido como SEL por su siglas en inglés³ (Sidorov et al., 2012), e incorpora una heurística para el tratamiento de negadores y modificadores.

El analizador de emociones a nivel de sesión calcula el perfil emocional general de la sesión y para cada diputado determina en qué porcentaje ha expresado alguna de las seis emociones de Ekman. En primer lugar, realiza un preprocesamiento sobre el texto asociado a cada diputado: i) tokenización mediante expresiones regulares, ii) extracción de raíces utilizando el paquete NLTK Snowball stemmer⁴, iii) eliminación de palabras vacías, y iv) conversión del texto a minúscula. En segundo lugar, realiza la clasificación de las emociones presentes en el texto de la siguiente forma:

```

1: diputado = {'ira': 0, 'asco': 0, 'miedo': 0,
2:   'alegría': 0, 'tristeza': 0, 'sorpresa': 0}
3: for palabra in texto do
4:   if palabra in lexicon then
5:     if anterior is negador then
6:       peso = 0
7:     else if anterior is increment then
8:       peso = 1,75
9:     else if anterior is atenuante then
10:      peso = 0,75
11:    else
12:      peso = 1
13:    end if
14:  end if
15:  for emocion in lexicon[palabra] do
16:    fpa = lexicon[palabra][emocion]
17:    diputado[emocion] += peso * fpa
18:  end for

```

El analizador de emociones a nivel de diputado calcula su perfil emocional en base a sus intervenciones en todas las sesiones y en qué porcentaje ha expresado cada una de las seis emociones a lo largo del tiempo.

3.3 Módulo 3: aplicación web para visualización

Una vez extraída, procesada y analizada la información, se desarrolló una aplicación web para mostrarla. Esta aplicación hace uso del

³El lexicón SEL está formado por 2.036 palabras en español que están asociadas con al menos una emoción básica (ira, asco, miedo, alegría, tristeza y sorpresa) a través del factor de probabilidad de uso afectivo (FPA por su siglas en inglés).

⁴http://www.nltk.org/_modules/nltk/stem/snowball.html

framework *Flask*⁵ de Python, de los lenguajes HTML, CSS y JavaScript, y de la herramienta Google Charts⁶.

Tal y como se ha comentado en el módulo 2, el analizador de emociones permite realizar dos tipos de análisis: uno a nivel de sesión y otro a nivel de diputado (Figura 3).



Figura 3: EmonCon - Página principal

Mediante la opción *sesión*, la aplicación web permite filtrar por fecha de sesión, consultar el PDF de cada sesión y visualizar el perfil emocional completo de la sesión (Figura 4) y de un diputado concreto ese día (Figura 5). En el perfil general de la sesión aparecerá un diagrama de sectores que indicará el porcentaje en el que ha estado presente cada emoción ese día y seis rankings, uno por cada emoción, en los que se mostrará los cinco diputados que más han sentido cada emoción. En el perfil del diputado seleccionado aparecerá un diagrama de sectores representando los porcentajes en los que ha expresado las distintas emociones en la sesión celebrada. Además, se mostrará una gráfica de barras en la que se podrá comparar las emociones del diputado ese día concreto con las emociones que suele mostrar, es decir, con su perfil habitual.

Mediante la opción *diputados* es posible seleccionar un miembro de los grupos parlamentarios (Figura 6), accediendo a su perfil emocional habitual por medio de un gráfico de sectores y de un histórico en el que se mostrará la evolución de este perfil del diputado (Figura 7). Además, en el diagrama del histórico se puede seleccionar un día concreto y acceder a la correspondiente sesión.

4 Conclusiones y trabajo futuro

En este trabajo se ha presentado EmoCon, un prototipo para analizar las emociones expresadas en el Congreso de los Diputados, con el objetivo de obtener y mostrar el perfil emo-

⁵<https://www.fullstackpython.com/flask.html>

⁶<https://developers.google.com/chart/>

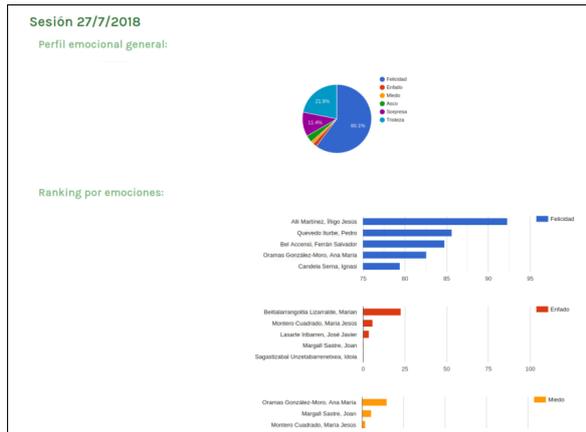


Figura 4: EmonCon - Análisis de una sesión: perfil completo de la sesión

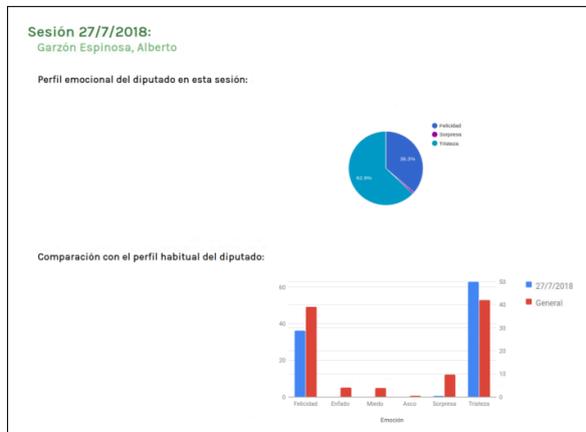


Figura 5: EmonCon - Análisis de una sesión: perfil de un diputado

cional a nivel de sesión parlamentaria y de cada diputado que interviene.

Como posible trabajo futuro de este sistema nos planteamos estudiar otros métodos de análisis de emociones y analizar la legislatura actual. Además, sería interesante añadir un apartado de temáticas para reflejar los temas tratados en cada sesión y así poder ver qué emociones se generan en función de lo que se esté hablando. Así mismo queremos analizar perfiles de usuarios y partidos políticos para comprobar si los miembros de un mismo partido tienen un perfil similar, analizando sus intervenciones en el Congreso y lo que publican en medios sociales.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el Fondo Europeo de Desarrollo Regional (FEDER) y el proyecto LIVING-LANG (RTI2018-094653-B-C21) del Gobierno de España.



Figura 6: EmonCon - Análisis de un diputado: selección del grupo parlamentario

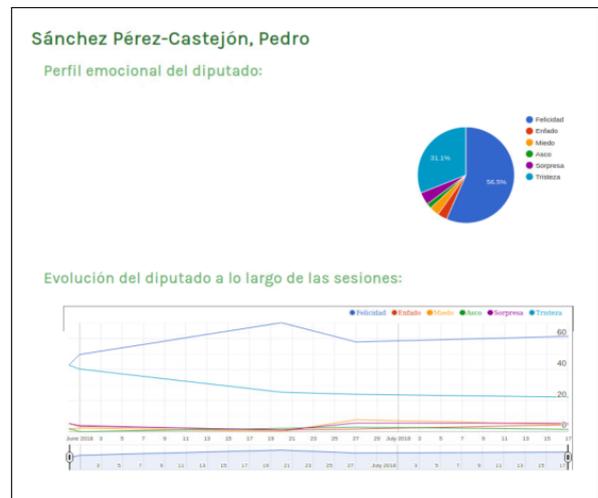


Figura 7: EmonCon - Análisis de un diputado: perfil general de un diputado

Bibliografía

- Cambria, E. 2016. Affective computing and sentiment analysis. *IEEE Intelligent Systems*, 31(2):102–107.
- Ekman, P. 1992. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200.
- Lövheim, H. 2012. A new three-dimensional model for emotions and monoamine neurotransmitters. *Medical Hypotheses*, 78(2):341 – 348.
- Plutchik, R., . K. H. 1986. *Emotion: Theory, Research and Experience*. New York: Academic Press.
- Sidorov, G., S. Miranda-Jiménez, F. Viveros-Jiménez, A. Gelbukh, N. Castro-Sánchez, F. Velásquez, I. Díaz-Rangel, S. Suárez-Guerra, A. Treviño, y J. Gordon. 2012. Empirical study of machine learning based approach for opinion mining in tweets. En *Mexican international conference on Artificial intelligence*, páginas 1–14. Springer.